

Speech Processing 2009/10

1st Test

April 15th 2009

Please identify this form with your name and student number in the reserved space at the bottom. The answers to multiple-choice questions will only be accepted if inserted in the appropriate place. Wrong answers will be penalized. The phonetic symbols should use the SAMPA alphabet (Lisbon accent).

1. Classify as True (T) or False (F)

- (a) Jitter is the period to period variation in amplitude of the vocal cords vibration.
- (b) Voiced plosives are characterized by a lower VOT than unvoiced plosives.
- (c) Assimilation is not very sensitive to speaking rate.
- (d) Frequency masking is asymmetric, i.e., a masker has more effect on tones above its frequency than on those below its frequency.
- (e) The zero crossing rate may be used alone for pitch detection.
- (f) The short-time average magnitude is preferred to the short-time energy for distinguishing between speech and non-speech segments.

2. Classify as True (T) or False (F)

- (a) Log-area ratios can be derived from reflection coefficients and vice-versa.
- (b) Error minimization in linear prediction corresponds to a better modeling of the formants than the valleys between formants.
- (c) The matrix of the covariance matrix is Toeplitz, allowing efficient recursive methods such as the Levinson-Durbin recursion.
- (d) The cepstrum of a pulse train only has non-zero values at integer multiples of the pitch period.
- (e) The cepstrum of a complex exponential $a^n u[n]$ decays more slowly than the signal itself.
- (f) The real cepstrum is the even part of the complex cepstrum.

3. Give examples of sounds for European Portuguese with the following properties (insert “-” in case they do not exist). Only one example is required for each category.

- (a) vowel high nasal
- (b) fricative voiced palatal
- (c) plosive coronal unvoiced
- (d) vibrant labial

4. Given that the maximum sensitivity of the human ear corresponds to the resonance of the auditory canal, compute the approximate distance from the external ear to the eardrum of a speaker with a maximum sensitivity at 3400 Hz. Indicate all steps of the computation.

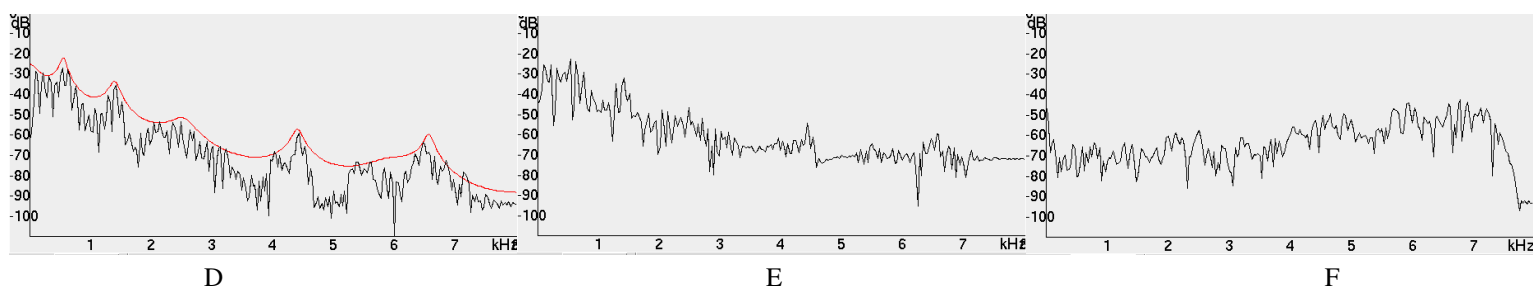
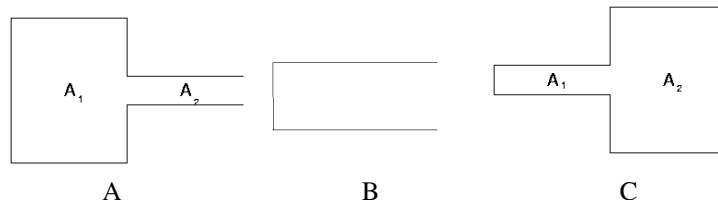
5. Complete the following sentences:

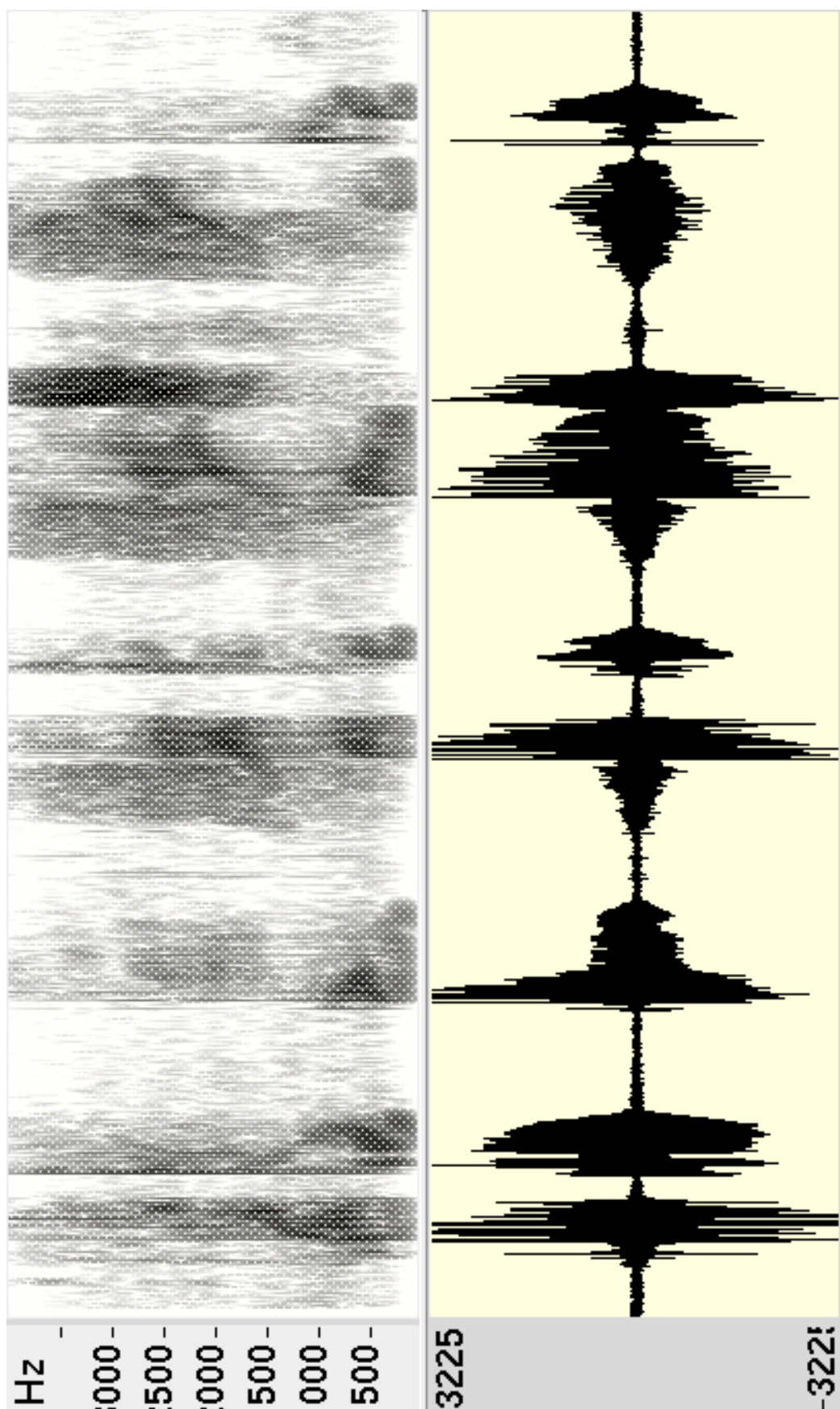
- (a) The MFCC representation uses a — scale, whereas the PLP representation uses a — scale.
- (b) Both MFCC and PLP representations emphasize — frequencies.

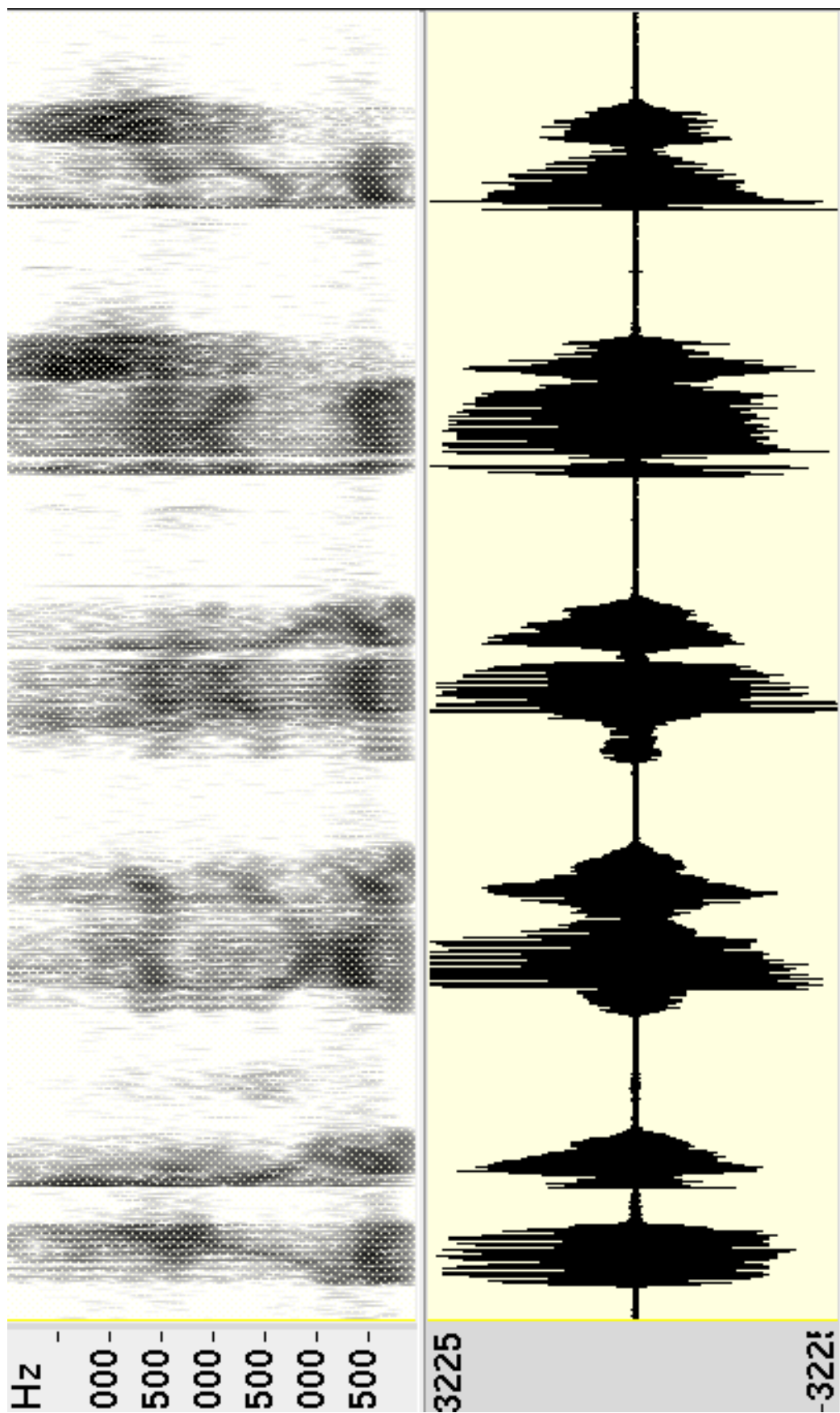
- (c) For intensity/loudness conversion, the MFCC representation uses a — compression, whereas the PLP representation uses a —.
6. Figures A, B and C correspond to vocal tract models of three vowels. Indicate the (European Portuguese) vowel for which each model is more appropriate.
 7. Write the expression for the linear prediction residual $e[n]$ of signal $x[n]$, both in the time domain and in the z-transform domain.
 8. Identify the 10-digit sequence by inspecting the spectrograms and waveforms in pages 2 and 3. There are no repeated digits. The recordings correspond to telephone speech.
 9. Figures D, E and F present the short-time Fourier transform (magnitude) of segments by the same speaker. The LPC spectral envelope is also shown in one of them.
 - (a) Indicate approximate values for F_0 , F_1 and F_2 (Hz) for figure D.
 - (b) To which of the 3 vowels of the extremes of the vowel triangle does it correspond?
 - (c) May it correspond to a window of 5 ms?
 - (d) May it correspond to a male voice?
 - (e) May the spectral envelope be obtained with an LPC analysis of order 8?
 - (f) Does it correspond to telephone speech?
 - (g) Figures D and E correspond to the same segment, but they have been obtained with a different window (either Hamming or rectangular). Please indicate which one was obtained with a rectangular window.
 - (h) Figure F corresponds to the residual of a vowel segment or to a fricative segment?
 10. Consider the following sentences:

Discriminação salarial leva enfermeiros a fazerem nova greve, a segunda em dois meses.
Plataformas sindicais esperam uma fortíssima adesão.

 - (a) Write the broad phonetic transcription.
 - (b) What are the missing liquids in this transcription?







Test 1 - Answers

Name:	
Number:	

1. (1.2 val.) Indicate T or F:

a	b	c	d	e	f

2. (1.2 val.) Indicate T or F:

a	b	c	d	e	f

3. (2.0 val.)

a	
b	
c	
d	

4 to 7. (2.0/1.0/1.2/1.4 val.)

4	
5 (a)	
5 (b)	
5 (c)	
6 (a)	
6 (b)	
6 (c)	
7	

8. (3 val.)

--	--	--	--	--	--	--	--	--	--

9. (4.6 val.) Complete:

F0	F1	F2	Vow	Y/N	Y/N	Y/N	Y/N	D/E	Res/fric

10. (2.0/0.4 val.)

(a)	
(b)	